Humor as a Tool (and a Weapon)

Introduction: Why Humor Matters

Humor is everywhere. But what exactly is humor, and why does it matter? While we often think of humor as laughter or mere amusement, but it plays a deeper role in shaping social interactions, reinforcing norms, and even challenging authority. Memes have become a staple in our everyday digital lives, serving not only as a source of amusement, but also as a way to bond over shared experiences—from last night's NFL game to political events and even large-scale crises like the Evergreen ship stuck in the Suez Canal. Viral videos, sitcoms, and stand-up comedy shape how we engage with humor and how we understand ourselves. But humor isn't always lighthearted. One of its oldest and most enduring forms is amusement at others' mistakes.

Take memes, for example. They aren't just funny pictures. They're ways of responding to crisis, flirting with someone, mocking power, or signaling who's in on the joke. What we find funny says something about who we are, and what we're willing to let slide. Playground bullies know this. So do comedians, politicians, and trolls.

This chapter argues that we should stop thinking of humor primarily as an emotional experience (amusement) and start thinking of it as a communicative *tool*: a kind of social technology. Rather than asking *what humor is* in some deep, metaphysical sense, we should ask *what humor does* in different communication contexts. This shift helps us make sense of humor's double edge: how it builds solidarity and excludes, empowers and silences, how it both resists power and reinforces it.

That might sound obvious to anyone who's ever been teased on the playground or found catharsis in a meme. But traditional philosophical theories of humor, focused on superiority, incongruity, or psychological relief, have mostly treated humor as a private state of amusement. These theories capture important features of the humorous experience, but they miss what humor is doing in actual social life: enforcing norms, dismissing concerns, inviting intimacy, or shielding bigotry.

In what follows, I survey the dominant philosophical theories of humor and show why they're inadequate on their own. I then develop an alternative approach that treats humor as a communicative practice with ethically significant effects, many of which are unintentional. By shifting focus from what humor "is" to what humor *does*, we gain a better grip on how to intervene when humor causes harm, and how to use it well.

Why Philosophers Care About Humor

Humor is a universal human phenomenon, shaped by culture much like language or morality. Because of this, philosophical inquiry into humor, laughter, and comedy connects with a range of disciplines, including ethics, philosophy of language, social philosophy, and philosophy of mind (to name just a few). Philosophers have been trying to pin down the fundamental nature of humor since at least Plato. Most agree that humor is best understood as the emotion of comic

amusement, but the real debate concerns what actually causes it, and whether there are limits to what *should*.[1]

Before diving into the traditional theories, we need to get clear on some basic terms. The "philosophy of humor" is a catchall for work on jokes, laughter, funniness, and the various practices that orbit comedy—like stand-up, satire, parody, and irony. But philosophers usually distinguish between three related domains: humor, comedy, and laughter.

Humor is the most general. It refers to the underlying phenomenon—not any specific joke or genre. Philosophers working in this space tend to ask foundational questions about humor's biological, psychological, or developmental origins. Why do babies laugh at peekaboo? What can joking about taboo topics reveal about how we process anxiety or disgust?

Comedy, by contrast, is more specific. It includes particular genres—slapstick, deadpan, cringe, surreal—and their social, aesthetic, and ethical dimensions. Why is *Nathan For You* so painfully funny? Is Borat's satire morally permissible if the people filmed didn't know it was a joke? What makes *Curb Your Enthusiasm* unbearable for some, cathartic for others?

Then there's laughter. Some philosophers, like Henri Bergson, have treated laughter itself as the central subject. For Bergson, laughter is a social correction mechanism—it kicks in when someone deviates from the script. More recently, theorists like Abrahams (2021) argue that humor just is any social practice aimed at eliciting laughter. But this raises questions of its own. Does that mean a dark joke that makes people wince rather than laugh fails as humor? And is laughter even under our control, or is it a reflex—like flinching, or doomscrolling?

These distinctions help clarify the terrain. But they also point to a deeper issue. While these theories focus on what makes something funny, I focus on what people *do* with funniness once it's out in the world. The classic accounts all treat humor as an *inner reaction*: comic amusement. We are amused by jokes, but we also use humor to flirt, to evade, to police the vibe. We forward memes not just because they're funny, but because they say something *for* us. A Spongebob Squarepants meme can function like a mood ring. So can a TikTok clip dropped in group chat, or a sarcastic "Congrats on being so *brave*" when your roommate finally takes out the trash.

These are all things humor *does*. And they often happen under the radar of traditional theories, which treat amusement as the main event. But once we widen our lens, the picture gets richer. Humor becomes a kind of *social currency*: sometimes stabilizing, sometimes subversive, often more than mere laughter.

The "Big Three" Philosophical Theories of Humor

A standard method in philosophy is conceptual analysis, which asks: What is X? In our case: What is humor? Most philosophers agree on one half of the answer—humor is, at its core, an internal experience. But the debate is over what *causes* or *justifies* this amusement. What makes something funny in the first place?

The three dominant philosophical accounts of humor, Superiority, Incongruity, and Relief Theory, attempt to answer that question by identifying the proper object of comic amusement: others' inferiority, the unexpected, or psychological release. Each theory captures something

intuitive. But none fully explain humor's messy, shifting role in real life. While these theories focus on what makes something funny, I focus on what people *do* with funniness once it's "out there" in the world.

The classic accounts treat humor as a kind of internal flash—a private response of comic amusement. But that framing misses what humor actually does in the world: it travels through shared spaces, reshapes what can be said, and shifts how responsibility and risk are distributed across a group.

Let's look at each theory in turn.

Superiority Theory

Open any social media app and scroll for 60 seconds. You'll likely find a video of someone tripping, failing, or generally being a mess. Entire genres of internet content—fail compilations, prank videos, reality TV bloopers—are structured around amusement at others' expense. Some people (myself included) find these videos almost medicinal. One solid pratfall montage can cure a depressive spiral. But why do we find others' misfortunes so funny?

According to Superiority Theory, humor is amusement at another's inferiority or misstep—what Thomas Hobbes called a feeling of "sudden glory" at our own relative safety or status. This is one of the oldest philosophical accounts of humor, with traces in Plato, Aristotle, the Bible, and early Islamic philosophers. Plato warned that laughter signals a loss of rational self-control (Morreall 2023). This helps explain why jokes often fail when they hit too close to home. If you recognize yourself in the punchline, it's hard to find a joke funny. This is also why some people find cringe humor unbearable: it reminds them *too much* of their own social failures. My grad advisor, for example, couldn't watch *Curb Your Enthusiasm*—the secondhand embarrassment was too intense. The Superiority Theory explains why: if you don't feel like the one *above* the cringe, you're not laughing.

But there's a darker side. Laughing at others can quickly become a tool of exclusion, shame, or domination. Trolls, for instance, thrive on false sincerity—they bait others into being earnest so they can mock them. The joke is always on whoever takes things too seriously. As Henri Bergson (1911) put it, laughter can function as a "momentary anesthesia of the heart".

That said, superiority isn't always aligned with cruelty. Sometimes it flips. Humor can be used to reverse power dynamics, especially by those who are presumed inferior. Comedian Alan Shain, who is disabled, argues that humor allows him to confront dominant narratives about dependency and care. As he puts it, "humour (sic) allows a direct attack on dominant approaches to disability that frames disabled people as dependent and in need of caretaking" (2013, 338). Rather than being a passive target of pity, Shain deliberately chose, "the vehicle of comedy and humour (sic) as a way of bonding with an audience in order to promote critical thinking and dialogue around the meaning of disability" (337).

And this isn't just about disability. From drag queens skewering gender roles to baristas joking about customers in group chats, asserting playful superiority is often a way to assert oneself, not just to mock others. In these contexts, humor functions as a strategy for expressing one's agency.

Of course, not all humor fits the Superiority Theory. Puns, for instance, are not funny because they demonstrate grammatical superiority. Many internet memes aren't funny because they're mocking someone, but because they juxtapose completely absurd images or ideas. This brings us to our next philosophical account of humor: the Incongruity Theory.

Incongruity Theory

Why is it funny when a Victorian painting is captioned, "MFW someone says 'let's circle back' in an email"? Or when a cat is Photoshopped into the Last Supper? The Incongruity Theory holds that humor arises when we encounter a mismatch between expectations and reality and experience a cognitive dissonance that's funny rather than frustrating.

This view has roots in 18th century German philosopher Immanuel Kant, who described humor as "the sudden transformation of a strained expectation into nothing." More recently, philosophers like Noël Carroll (2014, 2020) and Jerrold Levinson (2006) have expanded the incongruity theory to encompass a wide range of comedic cases, from knock-knock jokes to avant-garde sketch shows.

Work in developmental psychology and cognitive science draws on incongruity theory to support the idea that humor is an innate human capacity. Babies around ten months old begin to display laughter in response to incongruities—like when a familiar object is used in a surprising way. Evolutionary psychologists suggest this response might be adaptive: incongruity detection helps us catch errors or outliers in our environment without panicking. Humor becomes a kind of safe cognitive glitch. Evolutionary and cognitive scientists have proposed that humor functions as a *debugging mechanism*—helping humans detect and process errors in thinking or perception without experiencing negative emotional consequences.[2]

But like Superiority Theory, Incongruity Theory struggles to explain what humor *does*. The incongruous might amuse, but it can also confuse, alienate, or provoke rage. Think of rage bait content—someone calmly pouring milk on pizza, or putting a Tide pod in spaghetti. The incongruity doesn't amuse; it aggravates. These formats hijack the glitch in our expectations, not to make us laugh, but to make us *engage*.

And then there's care trolling: posts that feign concern in order to bait others into outrage or correction. It's incongruous because it pretends to care while clearly mocking being concerned, but its function is *strategic*. Trolling uses the audience's expectations against them. You "get" the joke only if you realize it's bait. If not, you risk *becoming* the joke. The success of care trolling depends on shared norms, stakes, and expectations, not just surprise.

Similarly, consider the court jester. On one level, the jester's humor relied on *incongruity*, presenting exaggerated foolishness to delight and surprise. On another level, it functioned *socially* as a mechanism for regulating authority—offering rulers a sanctioned space for critique while maintaining their dominance.[3] This dual role highlights how humor can work both as cognitive play and as a social tool for reinforcing or challenging power structures (which is why the jester's function is more fully explored in the discussion of Superiority Theory). If humor is

sometimes more about regulating behavior, correcting errors, or even maintaining political stability, then the Incongruity Theory alone isn't enough to explain its full function.
Finally, some forms of humor seem more about managing emotions than about recognizing incongruity. Who among us hasn't spent an evening roasting our friends in the group chat? Is this an example of Superiority Theory (amusement at one's relative status) or Incongruity Theory (enjoying the tension between insult and affection)? And this is where the incongruity model falters. While it can tell us why we find a toddler in a tuxedo hilarious, it doesn't tell us why we'd meme it. Or who that meme flatters, mocks, or excludes. For that, we need to look beyond amusement, toward what the humor *does* in different contexts.
Sometimes what makes a joke land isn't the mismatch between what we expected and what we got, it's the chance to convey what we couldn't otherwise say. The incongruity model helps explain how a joke surprises us, but not why we needed that surprise in the first place. In some cases, the joke isn't pointing out an error, it's releasing pressure. That's where Relief Theory comes in.

Relief Theory

Picture yourself at a tense dinner. Someone cracks a self-deprecating joke—"Well, if the planet burns, at least I won't have to grade finals." Laughter breaks out. Not because it's especially clever, but because it punctures the silence.

Relief Theory treats humor as a pressure valve, or a psychological coping mechanism. According to Sigmund Freud, jokes offer socially acceptable ways to express repressed thoughts, anxieties, or cultural taboos. In *Jokes and Their Relation to the Unconscious* (1905), Freud argued that we laugh when we finally let something slip—something that was already there, but had been held back. The logic of Relief Theory is clearest in so-called gallows humor: jokes about suffering, mortality, trauma, or grief. These jokes aren't funny *in spite* of the darkness, they're funny because they help us cope with it. In a later 1927 essay, *Humor*, Freud praised gallows humor as evidence of the resilience of the human spirit—suggesting that humor enables us to acknowledge harsh realities while emotionally distancing ourselves from them. If humor involves an "anesthesia of the heart" (Bergson), then it makes sense that we sometimes find amusement in things that would be offensive or objectionable in serious contexts.

This framing also appears in the work of Viktor Frankl a Holocaust survivor and author of *Man's Search for Meaning* (1946). Frankl emphasized humor's role in survival, and he described humor as a "weapon of the soul." While imprisoned in a Nazi concentration camp, Frankl observed that laughing at absurdities (such as roll call at 4 a.m.) was a way to hold onto humanity under dehumanizing conditions. Humor, in that context, wasn't just a distraction. It was an act of defiance and self-preservation.

Humor allows us to publicly acknowledge taboo ideas without fully endorsing them, which may explain why gallows humor is more common in high-stress environments. Social scientific research has found that professions like firefighters, veterans, surgeons, and morticians rely on gallows humor as a coping mechanism.[4] Emergency room doctors, for instance, may joke about tragic aspects of their work—not because they find tragedy funny, but because humor helps them manage the emotional burden of constant crisis. But gallows humor isn't the only place where humor serves as a form of relief. The comedy roast, for example, shows how humor can be both brutal and deeply affectionate.

Similarly, we use play-insults with close friends, not to re-establish superiority or express pent-up aggression, but to reinforce social bonds. If Relief Theory were the whole story, we might expect roasting to function primarily as an outlet for repressed hostility. Yet, people often *invite* being roasted by their friends—and if the popularity of the Reddit forum *r/RoastMe* is any indication, people also enjoy being roasted by strangers as well.

But the cause of the psychological relief associated with humor isn't always straightforward. Research suggests that exposure to dark humor doesn't always alleviate psychological strain. In some studies, repeated exposure to gallows humor correlates with **lower** overall well-being.[5] So maybe it's not the joke itself that heals; it's **the social connections that make the joke possible.**

Relief Theory also helps explain culturally specific humorous traditions, especially those that emerge in response to systemic oppression. **Black Laughter**, a tradition among Black Americans during and after chattel slavery, offers a particularly powerful example. **Trickster tales,** a staple of Black storytelling, often feature a smaller animal outsmarting a larger predator—a symbolic stand-in for enslaved people resisting white oppressors. On a Relief Theory reading, these tales offered emotional release from the horrors of daily life.

But as Glenda Carpio argues in *Laughing Fit to Kill* (2008), Black Laughter is more than a coping strategy—it's a cultural technology of resistance. Trickster tales don't just entertain or defuse stress; they communicate survival strategies, transmit cultural memory, and sustain collective identity. The laughter they elicit isn't merely relief—it's recognition, instruction, and endurance. More broadly, humor in marginalized communities often functions not just to *let something out*, but to *hold something together*.

This raises a broader point: even when humor does relieve tension, that relief is never free-floating, it depends on *who's speaking, who's laughing, and what the laughter is doing*. A joke can discharge anxiety, sure, but it also marks who belongs, what's permissible, and who can say what without consequence. In that sense, relief is never just a feeling, it's also always a *function* within a social context.

This also helps explain humor's mixed psychological effects. If the primary function of gallows humor were emotional relief, we'd expect it to improve well-being. But some research suggests the opposite: repeated exposure to dark humor is often associated with lower psychological

resilience. This isn't because the jokes "don't work", it's because the *context* and *relationships* in which humor circulates often matter more than the emotional punchline.

So perhaps the therapeutic effect isn't in the humor itself, but in what it signals. The moment of relief might not be "ah, I got that off my chest," but "you saw that too." Relief, in that case, isn't catharsis: it's *recognition*. Humor is not just about expressing what is repressed; it is also a communicative tool that helps shape relationships, regulate behavior, and challenge authority.

And that brings us back to the larger issue all three theories struggle with. Superiority, incongruity, and relief each offer insight into what makes us laugh, but none of them capture what humor *does* once it leaves the individual mind. Laughter may start in the head, but humor rarely stays there. The traditional theories treat humor as a psychological reaction: a flash of superiority, a glitch in expectations, a release of tension. But even when those reactions are real, they don't explain why a roast binds some and breaks others, or why the same meme can soothe one group and alienate another. They help us locate the moment of amusement, but not the consequences of the joke in our social world.

To move forward, we need to stop asking what humor is and start asking what it *does*. Humor is not just a feeling—it's a way of communicating, a way of managing social expectations, and a way of deciding who counts and whose concerns get dismissed.

The Ethics of Humor

If we take humor seriously as a communicative act—not just a source of amusement—then we also have to take seriously the question of ***responsibility***. That's where many of the traditional philosophical accounts of humor run into trouble.

When someone makes a joke that lands awkwardly—maybe it's offensive, maybe it stings, maybe no one knows how to respond, it can be hard to figure out what just happened. Did the speaker *mean* it? Was it *just for fun*? Are you *overreacting* if you don't laugh? Or *complicit* if you do?

Humor muddies the waters. Unlike serious speech, where we usually expect people to mean what they say, jokes often work *indirectly***.** They play with tone, exaggeration, implication, and shared context. That's part of what makes humor powerful—but also slippery. It's easy to say something risky or even harmful, and then walk it back with, "Relax, I was just joking."

One reason jokes can be so hard to pin down is that they don't need to be true or even believed by the person telling them. A joke can land, amuse, or offend without the jokester personally endorsing its content or intending for it to do so. You might laugh at a joke you don't agree with. Or tell one you don't fully understand. What matters is that it *hits*, not that it's literally *correct*. In that way, jokes resemble what Aristotle called enthymemes in his *Rhetoric*—arguments with something left unstated, something the audience is expected to supply on their own. The power of a joke often comes from what isn't said out loud, but is still understood. That's what makes humor socially efficient, but also ethically slippery. It lets something circulate without being fully claimed. And that's why humor complicates responsibility. When a joke offends or harms, it's easy to say, "I didn't mean it," or "You're taking it too seriously." But the joke has still done

its work, because anyone who "got it" filled in the missing piece. Whether or not the speaker "meant it," the message was *received*.

Philosophers have tried to make sense of this slipperiness in different ways. Some, like **Noël Carroll** (2014) and **Simon Critchley (2002),** argue that jokes function like fiction: we enter a space of imaginative play where literal truth doesn't apply. Others, like **Ted Cohen (1999),** suggest that jokes are inherently ***interactive*** because they depend on shared background assumptions and require the audience to do part of the work. This helps explain why jokes don't need to be endorsed to succeed, but also why they can't be brushed off as harmless either.

Once we see humor as a collaborative act—something that circulates socially, relies on shared inference, and shapes the conversational field—we can better understand the ethical questions it raises. Let's look at three of them.

1. ***Can Immoral Jokes Be Funny?***

This question is often framed as the **moralism vs. immoralism** debate in philosophy.[6]

- **Comic moralism** holds that a joke's immorality can count against its funniness. That is, laughing at a racist or sexist joke may be a kind of mistake—something we *should* not find funny.
- **Comic immoralism**, by contrast, insists that humor and morality occupy separate registers: a joke can be deeply offensive and still be funny. In fact, the offensiveness might be part of what makes it land.

Most philosophers take a moderate view: that immorality *can* undermine funniness, but doesn't always. But even this assumes that amusement is the central measure of a joke's value, and that the question is about **individual reactions.**

If we treat humor as a **social act,** the better question is: *what is this joke doing in the room?* Whose expectations does it reinforce? Who is positioned as the target? What kind of conversational space does it help create?

2. ***Can Jokes Harm? And If So, How?***

This is where questions about **harm** enter the picture. Jokes are often protected by the claim that they're "just jokes" or not meant to be taken *seriously*. But humor doesn't need to be serious to be effective. A joke can communicate a stereotype, reintroduce a harmful idea, or normalize an exclusion *without ever asserting it outright*. When audiences fill in the unstated elements to "get" a joke, they participate in reinforcing its content, even if they don't believe it. The harm may not be direct or intentional, but it's still *real*: it's in what the joke makes easier to say next time, or harder to challenge in the moment.

Philosopher **Luvell Anderson** (2015) argues that focusing too much on the speaker's attitude misses this point. A joke doesn't need to reflect someone's personal beliefs to do ethical damage. Its power lies in how it shapes **uptake**, invites **alignment**, and precludes objection.

3. ***What Does Accountability Look Like in Humor?***

This brings us to the question of **responsibility**. If jokes don't require belief, if harm can occur without intent, and if laughter is often involuntary—then what does accountability even mean in the context of humor? Is it the case that "anything goes"?

Of course, accountability can't just mean policing what comedians "really think," or blaming people for laughing in the moment. This would be to assume that a successful joke must be true or believed. But it also can't mean shrugging and saying "it's just a joke." Because jokes still enter the conversational record.[7] They shape what can and can't be said next. And they shape who feels validated, visible, or shut down.

That's why I argue that ***accountability in humor must be collective.*** It's not just about the joker, or even the initial audience. It's about what a joke licenses, how it circulates, and who steps in, or stays silent. A joke succeeds because the audience helps it succeed. That means everyone who gets it is *involved* in its communicative status, whether they laughed or not.

So, the ethics of humor isn't just about amusement or offense. It's about participation. It's about whether we recognize the role we play in letting a joke land—and whether we're willing to change the conditions that make that landing easy.

What Humor Does

So far, we've looked at how philosophers try to define humor and how they debate its ethical implications. But to understand humor's full impact, we need to change the question. Instead of asking *what humor is*, we should ask *what humor does*. Humor isn't just a feeling. It's something we *do*—a tool we use to manage relationships, reinforce norms, and influence what people take seriously. It shapes social space.

To get a better grip on humor's role in communication, we need to shift our focus. **Following pragmatist thinkers like John Dewey and Amie Thomasson**, we can think of humor not as a cognitive state, but as a *tool*—something we use to achieve different ends.[8] Consider a hammer. Its primary function is to pound nails. But it can also be used as a paperweight or, more nefariously, as a weapon. The same is true of humor. Humor's social function explains why it can be used in radically different ways—from fostering camaraderie to avoiding responsibility.

The shift from emotion to function helps us make sense of why humor is so ethically slippery. If humor were just amusement, then we could evaluate jokes the same way we evaluate other forms of entertainment: Did it make you laugh? Was it funny enough to be worth the risk? But humor isn't just about laughter. It's about interaction. We send memes not just because they're funny, but because they say something for us. We tease our friends not only to amuse ourselves, but to maintain closeness. We tell jokes to flirt, to bond, to distract, to deflect, to test the waters. And we laugh at things that are cruel, confusing, or even offensive—sometimes without meaning to, and sometimes because laughing is the easiest way to move on.

This is why I argue that we should treat humor as a communicative tool. Like language, humor allows us to do things—especially things that would be harder to do in straightforward speech. Humor often works where directness would fail, or has failed in the past. A caregiver might

jokingly ask a toddler, “What’s the magic word?” instead of commanding them to, “Say please.” It’s still a correction but softened through play.

Importantly, the psychological and communicative functions of humor don’t always line up. A person might make a joke because they think it’s funny, but the joke still *does something* whether or not that’s what they intended. Teasing a child might feel lighthearted, but it also teaches boundaries. Sharing a meme might just feel fun, but it reinforces who’s “in” on the joke, and who isn’t.

This distinction matters because it shows that humor isn’t just a reflection of private feelings—it’s a public act with social consequences. You don’t have to *intend* to exclude someone for a joke to function that way. Humor works, even when it works unintentionally.

Humor as Silencing. Thinking of humor as a tool also helps explain its unintended consequences. Jokes don’t just entertain, they can silence, dismiss, or delegitimize. Take this example: A young intern raises concerns about her boss, an influential movie producer, inviting young women into his office for “closed-door” meetings. A coworker smirks and says, “What, are you *jealous*?”—prompting the others to laugh. She drops the subject altogether. What just happened?

For 20th-century Mexican philosopher **Jorge Portilla (2021)**, this is an instance of *relajo*: an event in which humor displaces the seriousness of value and undermines the conditions for solidarity. The coworker’s joke doesn’t just mock the intern—it frames her concern as unserious. The coworkers’ laughter seals the deal, making it harder for her to bring it up again. The result? The joke functions as a form of **silencing**. Portilla saw *relajo* as a cultural phenomenon, one that uses levity to deflate moral seriousness and make critique feel out of place.

Humor as a Weapon. If humor can be used to dismiss legitimate concerns, it can also be a **communicative weapon**. Consider **Donald Trump’s use of humor in political discourse.** He regularly deployed derogatory nicknames (e.g., “Sleepy Joe,” “Shillary”) and shared inflammatory memes online. This strategic unseriousness allowed him to deflect criticism while rallying supporters. He shaped the conversation on his own terms, and when pushed, could always fall back on the familiar defense: *“It was just a joke.”*

Why was Trump so effective? Because **weaponized humor puts its targets at a rhetorical disadvantage**. Respond seriously, and you seem humorless. Push back, and you’re accused of overreacting. The joke does its work—and then walks away.

The Playground Bully. The mechanics here aren’t limited to politics. Imagine a kid on the playground being teased with: “Whatever you say, Stinky.” Other kids on the playground laugh. The target now faces a dilemma:

- Option A: “Actually I’m not stinky. I just showered.”
    - They take the joke literally and look humorless.
- Option B: “That’s not nice!”
    - They challenge the jokester’s intention, which the jokester can instantly deny by reminding them, “I’m just joking!”.

Either way, the jokester wins. They've created a rhetorical trap. The humor functions by making the insult disavowable. And this is one of humor's most effective—and dangerous—features: it can **corner** people. A joke can put its target in a rhetorical double bind. If they respond seriously, they look humorless. If they laugh along, they seem complicit. If they object, they risk becoming the new punchline. In this way, humor regulates the interaction while shielding the joker from responsibility. The more successfully the joke lands, the harder it is to push back without social cost.

To understand what's going on here, we need to look more closely at how jokes *actually* ***work.*** Philosopher **Ted Cohen (1999)** argued that successful jokes are both ***conditional*** and ***interactive.*** They're conditional because they depend on shared knowledge—cultural reference points, social context. They're interactive because the audience has to fill in the missing pieces for the joke to land. In this sense, Cohen's account of jokes is structurally similar to what Aristotle's account of **enthymemes**: arguments with key premises left unstated, which are then supplied by the audience.

Take a meme from early in the COVID-19 pandemic: *A picture of a locked-down city with the caption "Remember to be nice to your wives: the restaurants are closed!!"* The audience has to fill in the joke: that women are responsible for cooking, and if their husband were rude, they might ruin the dinner on purpose, and there are no other options because quarantine has trapped everyone at home. Even if the sharer doesn't *believe* these things, the joke *functions* as if they do. The audience "gets it" by accommodating the joke's unstated assumptions…by supplying the missing premises.

Presupposition Accommodation. This interactive structure of jokes is also what philosophers of language call "presupposition accommodation". A speaker often says something that depends on an unstated assumption—and the listener goes along with it. If someone says, *"There are no pickles left,"* we don't assume pickles have gone extinct—we infer they're talking about *this* fridge. The audience supplies the missing context. This kind of uptake is so routine we barely notice it.

Jokes often work the same way. To "get" the joke, you accommodate the premise—even if you wouldn't normally accept it. And that's what makes humor **ethically loaded.** It introduces content into the conversational record without requiring anyone to explicitly stand behind it. Feminist philosophers like **Jennifer Saul** and **Kate McGowan** have shown how this mechanism allows sexist or racist jokes to smuggle in oppressive ideas under the guise of play.[9] The speaker can disavow responsibility, but the joke has already circulated. And with each retelling, the assumptions get a little more familiar, a little more acceptable. The more often these jokes are told, and the more often they're laughed at, the more easily those norms circulate. And because challenging a joke often comes with social cost, most people let it pass. This creates what Saul (2017) calls a "poisoned communicative environment**"**—one where oppressive ideas gain traction not by direct argument, but by subtle repetition.

When a joke smuggles in an objectionable presupposition—say, that women are bad cooks, or that men are emotionally inept—it can flatten complexity before it even arises. The laugh comes

easily because the frame is familiar. But that familiarity is exactly what allows oppressive ideas to pass unexamined. The joke lands, not because it persuades, but because it's already been heard before. Humor, in this sense, isn't just slippery, it's *saturated*. It relies on uptake that feels automatic, which is part of what makes it so difficult to interrupt.

What the Tool Model Reveals

So now we have a fuller picture of humor's function as a communicative tool:

- It allows speakers to float ideas without full commitment.
- It recruits the audience's cooperation through uptake and inference.
- It shifts what can be said next—often without scrutiny.
- It creates social costs for objecting or opting out.

That's why humor is so hard to challenge. And that's also why it's so powerful.

By moving from a psychological to an **instrumental** model of humor, we gain better tools for assessing what jokes *do*—who they serve, who they silence, and how they move through social space. This model doesn't pretend to settle every debate about humor's value or limits. But it does give us a sharper sense of how humor operates, and why it's worth taking seriously.

This is how humor can shape what philosophers like David Lewis and Robert Stalnaker call the conversational record—what gets taken for granted in a given context. Once a joke enters the record, even implicitly, it changes the background conditions of the conversation. It doesn't matter whether anyone stated the premise directly. If the audience went along with it, the idea is now *in circulation*. It's been normalized.

And challenging it isn't easy. If you interrupt a joke to question its assumptions, you risk becoming the next joke. If you try to clarify what was said, you're told you didn't get it. If you point out harm, you're "killing the vibe." The social cost of interruption is often higher than the cost of silence…which is why silence usually wins.

Now we have fuller account of how humor functions as both a tool and a communicative weapon:

- It allows people to spread harmful ideas while maintaining plausible deniability.
- It normalizes stereotypes by embedding them in everyday social interactions.
- It makes challenging the joke socially costly (risking ostracization or harm).

Like any tool, humor can be used for good or bad. It can bond people together or push someone out. It can be a shield for the powerless, or a weapon for the already powerful. The same mechanism that makes a joke land with your friends can also make it impossible to challenge in public. So, we need better questions. Not just: *Was it funny?* Or: *Did they mean it?* But:

- What did the joke *allow*?
- What did it make harder to say *afterward*?
- Who was *protected*, and who was *exposed*?
- What vibe did it *set*, and who had to *adjust* to it?

When we ask what humor *does*, we start to see that jokes don't just reflect our social world, they actively help build it. And if we care about the kind of world we're building, we have to care about the jokes that lay its foundation. If humor creates social space, it can also reclaim it. The same tools that allow jokes to evade responsibility can be turned around to deflate authority, reframe assumptions, and shift the vibe entirely. That's the promise of counter-humor, not as a polite objection, but as a way of pushing back ***on humor's own terms.*** In what follows, I show how weaponized humor can be turned against itself, not by rejecting the game, but by flipping its rules.

Using Humor as a Weapon Against Weaponized Humor

However, all hope isn't lost. If humor can be used to smuggle harmful assumptions or stereotypes into the conversational record, it can also be used to interrupt them**.** The same mechanics that allow jokes to reinforce dominant norms can also be turned against them. Counter-humor isn't about being polite or offering an alternative—it's about responding to weaponized humor **on its own terms**, using the same tools to shift tone, deflate presuppositions, and reroute power. Humor is not only a communicative tool, it's a weapon, one that is particularly effective in contexts where literal speech might be punished, ignored, or dismissed. To counter an objectionable joke without escalating the conflict or incurring heavy social costs, the best move is often a **comeback**—a response that reshapes the joke's assumptions without directly challenging the joker. This might sound complicated, but in practice, it's intuitive. Consider a classic playground insult:

Bully: "You're so hairy, you make Chewbacca look naked."

What are the possible responses?

Option A: "I'm not *that* hairy!" (Responding literally, making oneself look defensive.)

Option B: "*That's* not nice!" (Challenging the bully's intentions, which invites further teasing.)

Option C: "Oh yeah? You should see my arms when I *don't* shave."

Option C is the winner, not because it "wins" the exchange, but because it plays within the rules to rewrite them. A witty comeback functions as a *deflection* rather than a direct *confrontation*, performing at least three communicative moves at once:

1. It *undermines* the insult's premise—the response reframes the joke, making the target appear unfazed.
2. It *shifts* social dynamics—rather than being positioned as a passive recipient of ridicule, the target takes control of the interaction.
3. It *forces* the audience to accommodate a different presupposition—instead of reinforcing the idea that excessive hair is undesirable, the joke normalizes it as an amusing non-issue.

This is more than clever repartee. It's a ***strategic communicative maneuver***. Comebacks like "I know you are, but what am I?" work because they disrupt the structure of the joke itself. Social

psychologists have found that even toddlers deflect teasing through play—suggesting this tool develops early and is widely understood.[10]

When humor is used to introduce harmful ideas, it does so in a way that distributes responsibility—no one person has to explicitly say, "*Sexist jokes are acceptable here*"; the joke implicitly establishes that norm for the group. When offensive humor goes unchallenged, it reinforces harmful norms and creates social conditions that make it easier for those norms to persist. But direct confrontation is often ineffective, as humor allows for easy deflection ("I was just joking"). This is why comebacks can be a powerful tool—they indirectly disrupt the joke's presuppositions without reinforcing them. However, crucially, the burden of countering weaponized humor cannot fall solely on the person being targeted.

Who Should Push Back? It's not fair, or realistic, to expect that the person being targeted should always be the one to deflect harmful humor. That's a structural burden, and one disproportionately placed on the most marginalized. A woman being catcalled, a Black student in a mostly white classroom, a trans person facing a transphobic meme online, each may be strategically *unable* to respond without risking social, physical, or professional harm.

Responsibility must be shared. Just as harmful humor relies on collective uptake to function, so does its undoing. A good comeback changes the air in the room—but so can a well-timed eye-roll, a derailing joke, or a shift in tone. Crucially, these interventions don't require the target to defend themselves. In fact, the most powerful counter-humor often comes from *bystanders*, especially those with more social protection. As Lisa McKeown (2022) argues, resisting humor's harmful effects isn't just about the speaker or the target—it's about the social space the joke creates and everyone who participates in it. McKeown shares a story where she was catcalled, and her guy friend redirected the joke as if it were aimed at him—diffusing tension and disarming the speaker without confrontation. That's the power of weaponized counter-humor: it operates within the same space, but reframes the rules of engagement.

This brings us back to Portilla's analysis of *relajo*. In the intern's case, the joke operated to dismiss her concern, and this was the collective fault of her coworkers. But a different joke, or even a moment of unexpected levity that *refused* to laugh along, could have undone that effect. It could have re-opened space for solidarity. Similarly, the catcall deflection worked because it disrupted the social space in which the offensive remark was allowed to pass unchallenged.

Thus, humor functions socially, and its responsibility is social too. Just as *relajo* can diffuse accountability to the point that no one individual can be held responsible, counter-humor can operate as a shared intervention. Just as humor can shape social reality, so can our responses to it, and that is a collective task.

A truly pragmatist approach to humor must take seriously not only what humor *does* but who is in a position to challenge it, and how responsibility should be distributed. If humor can be used to reinforce oppressive norms, it can also be used to resist them. But this isn't just a matter of

“fighting fire with fire”—it’s about using humor’s unique properties to interrupt, reroute, and reshape the space of communication. Counter-humor doesn’t work by opting out of the game. It works by **rewriting the rules.** And in non-ideal conditions, such as contexts where direct confrontation carries risk, that may be one of the few tools available that can deflate harm without inflaming conflict.

Conclusion: Why Humor as a Tool Matters

As we’ve seen, humor can build solidarity, but it can also exclude. It can provide relief, but it can also reinforce social hierarchies. It can highlight absurdities, but it can also obscure serious concerns. None of these effects are reducible to amusement alone, and they can arise regardless of the jokester’s intention. Treating humor as a communicative tool—rather than just a psychological phenomenon—helps us move beyond debates about intention and amusement. It allows us to focus instead on how humor actually works in interactions, in social norms, and in power dynamics. This instrumental model doesn’t just clarify why certain jokes feel wrong—it helps us understand why they’re *hard to challenge*, how they embed assumptions, and how they structure the space of what can be said.

The practical consequences of this shift in perspective are significant. When humor is used to dismiss concerns—as in Portilla’s analysis of *relajo*—the problem isn’t just that someone found something funny. It’s that humor was used to displace seriousness and undermine the conditions for solidarity. When humor is weaponized, as we saw with Trump’s strategic political unseriousness, the issue isn’t just about taste or offense, it’s about how humor reshapes the public record and forces opponents to play on tilted ground. And when objectionable jokes rely on presupposition accommodation, the problem isn’t reducible to whether they were funny or meant seriously. It’s that they normalize harmful assumptions, quietly recalibrating what can and cannot be challenged going forward.

This is why thinking of humor in instrumental terms, as something that does things, offers a more useful philosophical framework. It gives us better traction on humor’s social effects, and better tools for resisting its harmful uses.

Before we wrap up, it’s worth returning briefly to the traditional view. If humor is defined as comic amusement, then a joke’s wrongness seems to turn on the wrongness of the amusement it generates. But this approach raises endless follow-up questions: Is laughter a form of endorsement? Can I laugh and disagree at the same time? What distinguishes endorsing laughter from distancing laughter? These are not bad questions, but they’re not the most urgent ones either. In real-world contexts, people use humor to exclude, deflect, and dominate. And they also use humor to survive, to resist, to connect, and to reroute power. Understanding *those* moves requires a different philosophical approach—one grounded not in internal feelings, but in social interaction.

That’s why this chapter didn’t just try to map humor’s conceptual structure. It made a case for treating humor as a tool: one that helps explain why jokes matter so much, why some hurt more than others, and why responsibility must be understood as shared. We already hold ourselves and

each other accountable for what we find funny. That's not a betrayal of comedy—it's part of what it means to have a sense of humor. What we need now is a better framework for making sense of those stakes. A model that helps us ask not just *"Was that funny?"* but *"What did that joke do, and what can we do, now, in return?"*

I want to end with a general note. It's important that comedy, as a public institution, isn't unnecessarily restricted. Humor matters because it *can* push boundaries, test norms, and speak where serious language sometimes fails. At the same time, we already care about holding ourselves and each other accountable for what we find funny—that's part of what it means to have a sense of humor. The real challenge is figuring out how to navigate that accountability. Too often, these discussions get stalled because they're framed around a comedian's beliefs or intentions—questions that rarely get us anywhere. But what if we treated humor as a tool, with real social effects that often arise unintentionally? What kinds of conversations would that make possible? What new responsibilities—and new opportunities—might emerge?

Discussion Questions

1. Is "just joking" ever a legitimate defense? How can we tell when humor crosses a line, and who gets to decide where that line is?
2. Why is it often socially costly to challenge a joke, even when it's harmful? What does this say about how humor manages power in a group?
3. If a joke works because of a harmful stereotype, even if no one explicitly says it, who's responsible for what it communicates? Is it enough to say "they didn't mean it"?
4. If you laugh at a joke that crosses a line, are you responsible for what it reinforces, even if you didn't agree with it? Why or why not?
5. When is it helpful to respond to a joke with another joke? How can humor function as a tool for political disruption or solidarity?

Further Reading

Classical and Foundational Theories of Humor. These texts offer historical and canonical accounts of what humor is and what it's for.

- Plato, *Philebus* (on malicious laughter)
- Aristotle, *Rhetoric*, Book II (on wit and the enthymeme)
- Thomas Hobbes (1651), *Leviathan*, Pt 1, Ch. 6 ("Sudden glory")
- Henri Bergson (1911), *Laughter: An Essay on the Meaning of the Comic*
- Immanuel Kant (1790), *Critique of Judgment*, sec. 54

Contemporary Theories of Humor. Core works in philosophy and aesthetics that develop cognitive, emotional, and social models of humor.

- Ted Cohen (1999), *Jokes: Philosophical Thoughts on Joking Matters*
- Simon Critchley (2002), *On Humour*
- Jerrold Levinson (2006), "Philosophy of Humor," in *The Oxford Handbook of Aesthetics*

- Noël Carroll (2014), *Humour: A Very Short Introduction*

Humor, Speech Acts, and Social Harm. These readings explore how humor functions in conversation, distributes responsibility, and reinforces or resists norms.

- Mary Kate McGowan (2019), *Just Words: On Speech and Hidden Harm*
- Jennifer Saul (2017), "Dogwhistles and Figleaves: Linguistics, Pragmatics, and the Disguises of Racism"
- Noël Carroll (1996), "Moderate Moralism"
- Aaron Smuts (2006), "The Ethics of Humor: Can Your Sense of Humor Be Wrong?"
- Matthew Kieran (2003), "Art, Imagination, and the Cultivation of Morals"
- Luvell Anderson (2015), "Racist Humor"

Works Cited

Anderson, Luvell. 2015. "Racist Humor." *Philosophy Compass* 10 (8): 501–509. https://doi.org/10.1111/phc3.12237

Bennett, M. P., Zeller, J. M., Rosenburg, L., & McCann, J. 2003. "The Effect of Mirthful Laughter on Stress and Natural Killer Cell Activity." *Alternative Therapies in Health and Medicine*, 2, 38–45.

Bergson, Henri. 1911. *Laughter: An Essay on the Meaning of the Comic*. Translated by Cloudesley Brereton and Fred Rothwell. London: Macmillan.

Carpio, Glenda R. 2008. *Laughing Fit to Kill: Black Humor in the Fictions of Slavery*. Oxford: Oxford University Press.

Carroll, Noël. 2014. *Humour: A Very Short Introduction*. Oxford: Oxford University Press.

Carroll, Noël. 2020. "Comic Amusement, Emotion, and Cognition." *The Monist* 103 (2): 205–228. https://doi.org/10.1093/monist/onaa004

Cohen, Ted. 1999. *Jokes: Philosophical Thoughts on Joking Matters*. Chicago: University of Chicago Press.

Critchley, Simon. 2002. *On Humour*. London: Routledge.

Freud, Sigmund. (1905) 1963. *Jokes and Their Relation to the Unconscious*. Translated by James Strachey. New York: W.W. Norton.

Freud, Sigmund. 1927. "Humor." In *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, Vol. XXI (1927–1931), translated by James Strachey, 159–166. London: Hogarth Press.

Hobbes, Thomas. (1651) 1981. *Leviathan*. London: Penguin Classics.

Hurley, M., Dennett, D., & Adams, R., Jr. (2011). *Inside Jokes: Using Humor to Reverse Engineer the Mind*. MIT Press.

Kianpour, Connor K. 2023. "Strong Comic Immoralism." *Journal of Aesthetics and Art Criticism*. 81 (3):363-377.

Levinson, Jerrold. 2006. “Philosophy of Humor.” In *The Oxford Handbook of Aesthetics*, edited by Jerrold Levinson, 429–450. Oxford: Oxford University Press.
Lewis, David. 1979. “Scorekeeping in a Language Game.” *Journal of Philosophical Logic* 8 (1): 339–359.
McGowan, Mary Kate. 2019. *Just Words: On Speech and Hidden Harm*. Oxford: Oxford University Press.
McKeown, Lisa. 2022. “The Joke’s On Who: The Performative Possibilities of Humor.” *Nordic Wittgenstein Review* 11 (1): 103–113. https://doi.org/10.15845/nwr.v11i1.3769
Morreall, John. 2023. “Philosophy of Humor.” *The Stanford Encyclopedia of Philosophy*, Winter 2023 Edition. Edited by Edward N. Zalta.
https://plato.stanford.edu/archives/win2023/entries/humor/
Otto, Beatrice K. 2001. *Fools Are Everywhere: The Court Jester Around the World*. Chicago: University of Chicago Press.
Portilla, Jorge. 2021. *The Suspension of Seriousness: On the Phenomenology of Jorge Portilla*. Edited and translated by Carlos Alberto Sánchez. Albany: SUNY Press.
Saul, Jennifer M. 2017. “Dogwhistles and Figleaves: Linguistics, Pragmatics, and the Disguises of Racism.” In *Implicit Bias and Philosophy, Volume 1: Metaphysics and Epistemology*, edited by Michael Brownstein and Jennifer Saul, 83–108. Oxford: Oxford University Press.
Shain, Alan. 2013. “Perspectives on Comedy and Performance as Radical Disability Activism.” *Journal of Literary & Cultural Disability Studies* 7 (3): 337–346.
https://doi.org/10.3828/jlcds.2013.28
Silva, Kathryn, and Barbara Rogoff. 2021. “Instructional Ribbing as a Cultural Practice for Guiding Children.” *Human Development* 64: 172–190. https://doi.org/10.1159/000517189
Thomasson, Amie L. 2025. *Rethinking Metaphysics*. Oxford: Oxford University Press.
Watson, K. 2011. Gallows Humor in Medicine. *Hastings Center Report*, 41(5), 37-45.

---

[1] The philosophy of humor has been explored across a range of traditions, from classical Indian and Chinese thought to Continental and pragmatic approaches. This chapter focuses primarily on debates within the analytic, Anglophone tradition, which has shaped much of contemporary philosophical inquiry into humor, particularly in terms of defining its structure, function, and ethical implications. At the same time, keeping these limitations in view lets us ask what gets left out and how other perspectives might push back on or complicate these mainstream accounts.

[2] Hurley et al., 2011.

[3] Beatrice Otto (2001) *Fools Are Everywhere: The Court Jester Around the World*

[4] For example. Watson (2011) on gallows humor in medicine.

[5] Bennett (2003).

[6] See Kianpour (2023) for more on the moralism-immoralism debate.

[7] The “conversational record” is a term from philosophy of language and pragmatics. It refers to the shared set of assumptions, topics, norms, and background knowledge that are treated as active in a conversation, even if no one states them outright. Once a joke introduces a stereotype or attitude into that shared space, it becomes part of what others take for granted going forward. See also David Lewis (1979), “Scorekeeping in a Language Game”.

[8] Thomasson (2017) argues that we can clarify metaphysical and conceptual questions by looking at what we do with our concepts—how they function in practice. This “instrumental” turn can help illuminate humor’s social and ethical stakes as well.

[9] Jennifer Saul (2017), “Dogwhistles, Political Manipulation, and Philosophy of Language” and Mary Kate McGowan (2019) *Just Words: On Speech and Hidden Harm*

[10] Silva and Rogoff (2021) found that toddlers can engage in playful deflection as a way of managing social interactions.